

# APLIKASI ANALISIS TABULASI SILANG (*CROSSTAB*) DALAM PERENCANAAN WILAYAH DAN KOTA

Imam Indratno dan Rahmat Irwinsyah

## ABSTRACT

*The field of quantitative planning analysis concerns with collecting, organizing and interpreting data to support planning processes, including analysis of category data. It is an appropriate methods to analyze discrete data. This analysis, also called as tabulation analysis, is used to determine whether the variables are related to one another or independent, and to measure the extent of their association. To use this cross tabulation analysis, the types of measurement data and the types of categorical data should be understood. This analysis is carried out by determining if the observation frequencies of occurrence of the categorical values of a qualitative could allow us to reject a hypothesis the expected frequencies of occurrence. Chi-square statistics is used to determine whether the observation frequencies differ significantly from the expected frequencies of occurrence. If computed  $\chi^2$  value greater than the value of  $\chi^2$  random variable with  $(r-1)(c-1)df$ , then the hypothesis of independent could be rejected. Furthermore, to determine or measure the extent of association, we use coefficient contingency with value of 0-1.*

## I. PENDAHULUAN

Kegiatan perencanaan kerap menuntut perencana untuk mampu menganalisis keterkaitan antar variabel atau faktor. Pengujian terhadap sejumlah variabel dibutuhkan dalam pengambilan keputusan untuk melihat keterhubungan maupun derajat keterhubungan (*asosiasi*) antar variabel atau faktor. Untuk itu setiap pengambilan keputusan harus didukung oleh informasi yang lengkap dengan melibatkan banyak pertimbangan dan kriteria maupun dukungan alat analisis sebagai alat bantu perencanaan.

Perencana dalam melihat obyek selalu dikaitkan dengan aspek keruangan. Pemahaman akan karakteristik wilayah secara menyeluruh (semua aspek) sangat dituntut bagi perencana. Pemahaman aspek keruangan harus sampai pada sub-sub wilayah yang akan menjadi dasar untuk menentukan dan memutuskan tindakan perencanaan yang paling sesuai pada keseluruhan wilayah yang direncanakan (Dewi Sawitri & Iwan P. Kusumantoro, 1992).

Pemahaman karakteristik wilayah yang mendalam dengan berbagai pertimbangan (variabel/faktor) yang dilibatkan menuntut perencana untuk mampu menganalisis keterhubungan antar variabel/faktor yang mempengaruhi perencanaan wilayah dan kota. Untuk melihat hubungan antar beberapa variabel/faktor akan lebih mudah apabila hubungan tersebut diterjemahkan dalam angka matematis (dikuantifikasikan). Pembaca akan mudah berbeda persepsi bila tingkat hubungan hanya diterjemahkan bukan dalam angka matematis (kualitatif).

Untuk melihat hubungan antar variabel atau faktor bergantung pada jenis data yang digunakan dalam perencanaan wilayah dan kota. Bila data yang digunakan berasal dari data sekunder maka analisis korelasi yang akan digunakan. Apabila informasi diperoleh dari data primer (*kuisioner*) maka analisis tabulasi silang (*cross-tab*) lebih cocok untuk digunakan.

Analisis tabulasi silang merupakan suatu prosedur dalam uji statistik untuk melihat

hubungan antar variabel atau faktor sekaligus memperoleh besarnya derajat keterhubungan atau asosiasi antar variabel atau faktor yang diukur. Pada bagian-bagian selanjutnya dari tulisan ini akan diperkenalkan metode tabulasi silang secara garis besar, meliputi sekilas analisis tabulasi silang, penggunaan, masukan data, keluaran data dan bagaimana menginterpretasikannya.

## II. SEKILAS ANALISIS TABULASI SILANG

Tabulasi silang merupakan metode analisis kategori data yang menggunakan data nominal, ordinal, interval, serta kombinasi di antaranya. Prosedur tabulasi silang digunakan untuk menghitung banyaknya kasus yang mempunyai kombinasi nilai-nilai yang berbeda dari dua variabel dan menghitung harga-harga statistik beserta ujinya. Contoh: ingin diketahui hubungan dan tingkat hubungan antara tipe rumah dan besarnya pendapatan.

Metode tabulasi silang (*crosstab* atau *cross classified*) memiliki beberapa metode pendekatan yang berbeda dan menggunakan uji statistik yang berbeda pula, bergantung pada banyaknya variabel yang akan diidentifikasi hubungannya satu sama lain. Jika hanya menggunakan dua variabel maka dapat menggunakan metode tabel kontingensi. Metode ini merupakan metode yang paling umum digunakan dalam analisis tabulasi silang. Jika variabel yang hendak diuji jumlahnya lebih dari dua, dapat menggunakan model yang disebut dengan *Hierarchical Log Linear*.

Tabulasi silang merupakan metode yang mentabulasikan beberapa variabel yang berbeda ke dalam suatu matriks yang hasilnya disajikan dalam suatu tabel dengan variabel yang tersusun dalam baris dan kolom. Variabel ini merupakan variabel kategori bebas pada satu bagian dan variabel kategori prediktor pada bagian lainnya. Tabel ini menunjukkan hubungan bivariat

antara pengukuran ketergantungan pada setiap variabel prediktor yang terpisah.

Analisis tabel kontingensi dua dimensi terfokus pada penentuan saling ketergantungan antara dua variabel. Diasumsikan terdapat  $N$  buah ukuran sampel,  $X_{ij}$  merupakan jumlah responden dari sampel pada setiap sel,  $X_{+i}$  merupakan jumlah margin baris dan  $X_{+j}$  merupakan jumlah margin kolom,  $i = 1, 2, \dots, j = 1, 2, \dots$ . Dikaitkan dengan tabel jumlah (*count*), tabel kemungkinan multi nominal diperoleh dengan cara membagi masing-masing sel dengan total ukuran sampel  $N$ .  $P_{ij}$  merupakan kemungkinan pada setiap sel  $(i, j)$  (Tabel 1).

Tabel 1

	$B_1$	$B_2$	Total
$A_1$	$P_{11}$	$P_{12}$	$P_{1+}$
$A_2$	$P_{21}$	$P_{22}$	$P_{2+}$
Total	$P_{+1}$	$P_{+2}$	$P_{++} = 1$

A dan B adalah dua variabel. Jika kategori saling bebas maka :

$$\frac{P_{11}}{P_{+1}} = \frac{P_{12}}{P_{+2}} = \frac{P_{1+}}{P_{++}} = \frac{P_{+1}}{P_{++}}$$

$$P_{12} = \frac{P_{1+}P_{+2}}{P_{++}}$$

$$P_{11} = \frac{P_{1+}P_{+1}}{P_{++}}$$

secara umum

$$P_{ij} = P\{\text{kategori baris} = i, \text{kategori kolom} = j\}$$

$$= P\{\text{kategori baris} = i\}P\{\text{kategori kolom} = j\}$$

$$P_{i+}P_{+j}, i = 1, 2, \dots, r, j = 1, 2, \dots, c$$

Dari distribusi multinomial dapat dilihat bahwa jumlah harapan dari observasi ketergantungan dalam sel  $(i, j)$  adalah  $m_{ij} = N P_{i+}P_{+j}$ . Margin observasi dapat digunakan untuk memperkirakan jumlah sel harapan dengan mensubstitusikan  $\chi_{i+}/N$  untuk  $P$  dan  $\chi_{+j}/N$  untuk  $P$ . Uji statistik ketergantungan dapat menggunakan *Statistik Pearson* atau uji Khi Kuadrat Pearson (*Pearson Chi-square test*). Uji Khi kuadrat ini adalah statistik yang biasa digunakan untuk menguji apakah ada hubungan an-

tara variabel baris dengan variabel kolom dalam suatu tabulasi silang.

$$\chi^2 = \sum_{ij} \frac{(\chi_{ij} - \chi_{i+} \chi_{+j} / N)^2}{\chi_{i+} \chi_{+j} / N}$$

Bila dalam perhitungan nilai  $\chi^2$  lebih besar dari nilai  $\chi^2$  variabel acak dengan derajat kebebasan  $(r-1)(c-1)$  maka hipotesis dua variabel saling bebas ditolak atau dengan kata lain dua variabel saling berhubungan.

Jika ditulis dalam bentuk lain, maka dapat dijelaskan sebagai berikut:

Bila N adalah ukuran sampel (banyaknya seluruh *case* yang diobservasi), maka probabilitas independen (P) dari suatu observasi yang berada dalam sel *ij* (baris ke-*i* dan kolom ke-*j*) adalah:

$$P_{ij} = \frac{(\text{jumlah baris ke-}i) \times (\text{jumlah kolom ke-}j)}{N}$$

sedangkan harga pengharapan (*Eij*) dari observasi dalam sel *ij* adalah:

$$E_{ij} = \frac{(\text{jumlah baris ke-}i \times \text{jumlah kolom ke-}j)}{N}$$

Statistik Chi-Kuadrat Pearson memenuhi persamaan sebagai berikut :

$$\chi^2 = \sum_{ij} \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

Keterangan :

*Oij* adalah hasil observasi (nilai frekuensi/count) pada baris ke-*i* dan kolom ke-*j*.

*Eij* adalah harga pengharapan (Exp. Val.) pada baris ke-*i* dan kolom ke-*j*.

### III. KEGUNAAN ANALISIS TABULASI SILANG

Analisis tabulasi silang bermanfaat dalam menyelesaikan permasalahan analisis data. Manfaat utama yang dapat diperoleh dari

analisis tabulasi silang, khususnya dalam perencanaan wilayah dan kota adalah:

1. Membantu menyelesaikan penelitian yang berkaitan dengan penentuan hubungan antar variabel atau faktor yang diperoleh dari data kualitatif. Penentuan hubungan akan digunakan sebagai dasar untuk penentuan tindakan perencanaan yang tepat pada tahap selanjutnya. Misalnya dalam menentukan kinerja pembangunan rumah sangat sederhana (RSS), salah satu faktor pertimbangan perencanaan adalah menentukan hubungan antara pengambilan tipe rumah dengan besarnya pendapatan. Dengan analisis tabulasi silang berdasarkan pada data primer yang telah diperoleh didapat suatu uji statistik yang akan diperoleh kesimpulan mengenai hubungan antara pengambilan tipe rumah dengan tingkat pendapatan.
2. Bila telah didapat hubungan antar variabel/faktor keuntungan, kegunaan kedua adalah dapat menentukan besarnya derajat asosiasi antar variabel/faktor tersebut. Dari angka yang diperoleh dapat diketahui tingkat keterkaitan antara pengambilan tipe rumah dengan tingkat pendapatan. Tingkat keterkaitan ditunjukkan dengan bilangan dari 0 sampai 1 yang menunjukkan derajat keterkaitan rendah hingga sangat erat.
3. Dapat menentukan variabel *dependent* (terikat) dan variabel *independent* (bebas) dari dua variabel yang dianalisis. Salah satu keluaran dari analisis tabulasi silang adalah penentuan variabel bebas dan variabel terikat yang ditunjukkan oleh *statistik lambda* atau *eta*.

Dapat dilihat bahwa analisis tabulasi silang akan sangat membantu perencana dalam menganalisis pada tahap selanjutnya, sebagai dasar dalam pengambilan keputusan. Analisis tabulasi silang berguna apabila data yang diperoleh merupakan data dalam bentuk data kategori yang diperoleh dari survei primer di lapangan.

#### IV. MASUKAN ANALISIS TABULASI SILANG

Untuk memasukkan data dalam analisis tabulasi silang, yang harus dipahami pertama kali adalah jenis/tipe data dan kategori data. Data yang berhubungan dengan teknik pengukuran atau analisis dikelompokkan menjadi 4 (empat) jenis data, yaitu:

##### Data Nominal

Data yang hanya menunjukkan kategori atau kelas, artinya data yang satu tidak dapat dibandingkan dengan data lainnya. Hubungan data nominal dapat dilihat melalui:

- jumlah kasus;
- mode (mean, median, kuartil);
- contingency correlation (Fisher's test, chi-square).

##### Data Ordinal

Data ordinal dapat dibandingkan dengan data lainnya, artinya data dapat dikatakan lebih baik, sama, atau lebih buruk dari data lainnya. Uji statistik yang digunakan antara lain :

- menghitung persentil data;
- median;
- korelasi ranking;
- sign test;
- run test.

##### Interval

Data dalam bentuk interval uji statistik yang dapat digunakan adalah :

- rata-rata aritmetik;
- standar deviasi;
- korelasi *product moment*;
- korelasi;
- t-test, F-test.

##### Metrik (rasio)

Data dalam bentuk rasio uji statistik yang dapat dilakukan adalah :

- koefisien korelasi;
- rata-rata geometrik;
- rata-rata harmonis.

Adapun tipe data yang digunakan dalam analisis tabulasi silang adalah tipe data ka-

tegori. Tipe data kategori ditunjukkan dalam tiga tipe tabel yang berbeda yaitu : tabel kontingensi, tabel ranking dan tabel multi dimensi.

##### a. Tabel Kontingensi

Tabel kontingensi merupakan cara menunjukkan variabel kategori yang mentabulasi silangkan antara satu variabel kategori dengan variabel lainnya. Hasil tabulasi silang ditunjukkan dengan tabel kontingensi (Tabel 2). Angka pada setiap tabel disebut dengan jumlah sel (*cell count*) atau frekuensi sel. Jumlah sel bila dibagi dengan seluruh sampel akan diperoleh frekuensi relatif kejadian dari kombinasi kemungkinan dari responden untuk dua variabel.

Tabel 2: Tabel Kontingensi 2 x 2

Menggunakan Produk Y	Menggunakan Produk X		
	Ya	Tidak	Total
Ya	82	10	92
Tidak	5	50	55
Total	87	60	147

##### b. Tabel Ranking

Tipe contoh tabel ranking diperoleh dari pertanyaan subyek sebanyak N buah untuk n obyek ranking yang disesuaikan dengan karakteristiknya. Misalnya data diperoleh dari pertanyaan pada 10 (sepuluh) ibu rumah tangga secara ranking pada penggunaan tiga jenis pasta gigi (Tabel 3). Tabel ranking dapat ditunjukkan dalam bentuk lain yaitu tabel *object-by-rank* (Tabel 4).

Tabel 3: Tabel Ranking

Rumah Tangga	Pasta Gigi		
	A	B	C
1	2	1	3
2	1	2	3
3	1	3	2
4	2	1	3
5	1	2	3
6	3	1	2
7	1	2	3
8	2	1	3
9	1	3	2
10	1	3	2

**Tabel 4:** Tabel Rank-Order (Nxn) dan Ranking-Obyek

Obyek	Ranking			Total
	1	2	3	
A	6	3	1	10
B	4	3	3	10
C	0	4	6	10
<b>Total</b>	10	10	10	

c. Tabel Multidimensi

Tabel kontingensi multidimensi merupakan hasil tabulasi silang dari tiga atau lebih variabel kategori. **Tabel 5** merupakan hasil tabulasi silang antara keinginan, status pernikahan, pendapatan dan jenis pekerjaan.

Dari berbagai jenis data tersebut di atas, analisis tabulasi silang merupakan analisis yang menggunakan masukan data dalam bentuk kategori. Oleh karena itu analisis tabulasi silang sering disebut juga analisis kategori data. Data kategori yang digunakan sebagai masukan terdiri dari empat jenis data, yaitu : data nominal, data ordinal, data interval dan data metrik/rasio.

**Keluaran Analisis Tabulasi Silang**

Keluaran analisis tabulasi silang terdiri dari :

- uji *chi square* untuk menentukan keterhubungan antar variabel atau faktor;
- pengukuran langsung yang terdiri atas pengukuran nominal dengan nominal, ordinal dengan ordinal dan nominal dengan interval;

- pengukuran kesimetrisan yaitu untuk menentukan variabel *dependent* dan *independent*.

Untuk melihat hasil keluaran analisis tabulasi silang dapat dilihat pada contoh penerapan sebagai berikut:

Pada suatu survei primer, hendak diketahui hubungan antara tipe rumah dan besarnya pendapatan keluarga. Untuk meneliti, telah ditetapkan variabel tipe rumah dengan tiga kategori yaitu: Tipe 21, 36 dan 45; dan tiga kelompok pendapatan keluarga yaitu: <Rp. 250.000,-, Rp. 250.000-Rp. 500.000,- dan >Rp. 500.000,-

Setelah dikompilasi, diperoleh kondisi data sebagai berikut :

Tipe Rumah	Besar Pendapatan	Jumlah Responden
1	1	15
2	1	8
3	1	6
1	2	6
2	2	10
3	2	20
1	3	5
2	3	18
3	3	12
	<b>Total</b>	<b>100</b>

Hasil tabulasi silang terlihat pada **Gambar 1**.

**Tabel 5:** Tabel Multidimensi Membeli x Status Nikah x Pendapatan x Pekerjaan

Pekerjaan	Pendapatan	Maksud Membeli			
		Ya		Tidak	
		Status Nikah		Status Nikah	
		Nikah	Belum	Nikah	Belum
A	< 10.000	10	22	2	4
	10 - 20.000	8	12	8	6
	> 20.000	11	5	12	20
B	< 10.000	13	7	12	8
	10 - 20.000	18	5	1	1
	> 20.000	42	42	5	2
C	< 10.000	12	8	15	6
	10 - 20.000	5	14	25	10
	> 20.000	15	5	13	7

**Gambar 1.** Tabulasi silang jumlah pendapatan dan tipe rumah

JUMLAH		TIPERUMA			Page 1 of 1
		Count	Exp Val	Row Pct	Col Pct
		Residual	Residual	Residual	Residual
		1.00	2.00	3.00	Total
< Rp. 250.000	1.00	1	0	5	6
	1.4	1.0	3.6	24.0%	
	16.7%	.0%	83.3%		
	16.7%	.0%	33.3%		
Rp. 250.000 - Rp.	2.00	2	2	6	10
	2.4	1.6	6.0	40.0%	
	20.0%	20.0%	60.0%		
	33.3%	50.0%	40.0%		
> Rp. 500.000	3.00	3	2	4	9
	2.2	1.4	5.4	36.0%	
	33.3%	22.2%	44.4%		
	50.0%	50.0%	26.7%		
Column Total		6	4	25	
		24.0%	16.0%	100.0%	

Number of Missing Observation: 0

Banyaknya *case* untuk masing-masing kombinasi nilai-nilai dari dua variabel ditampilkan dalam sel tabel bersama-sama dengan berbagai variasi persentase. Sel-sel ini menyediakan informasi mengenai keterhubungan antara kedua variabel tersebut. Pada tabel, **TIPERUMA** disebut sebagai variabel kolom (*column variable*), sedangkan **JUMLAH** disebut sebagai variabel baris (*row variable*).

Pada **gambar 1** tersebut, tiap sel berisi 5 baris informasi, yaitu :

- Baris pertama adalah jumlah *case* atau frekuensi dalam sel (*Count*). Contoh: 5 adalah jumlah *case* yang mempunyai jumlah pendapatan < Rp. 250.000 dan rumah tipe 45.

- Baris kedua adalah nilai harapan atau jumlah *case* yang diharapkan dalam sel tersebut (*Exp.Val.*). Misalnya: 1.4 adalah nilai harapan dari rumah tipe 21 dengan jumlah pendapatan < Rp. 250.000,-
- Baris ketiga adalah persentase baris (*Row Pct.*). Persentase baris ini diperoleh dengan membagi *Count* dalam suatu sel dengan nilai marjinal baris sel tersebut (*Row Total*). Contoh: 16.7% adalah persentase baris dari rumah tipe 21 dengan jumlah pendapatan < Rp. 250.000,- dan diperoleh dengan membagi 1 dengan 6.
- Baris keempat adalah persentase kolom (*Col. Pct.*). Persentase kolom ini diperoleh dengan membagi *Count* suatu

sel dengan nilai marjinal kolom sel tersebut (*Column Total*). Contoh : 33.3% adalah persentase kolom dari pendapatan < Rp. 250.000 dan rumah tipe 45 dan diperoleh dengan membagi 5 dengan 15.

- Baris kelima adalah persentase total (Tot. Pct). Diperoleh dengan membagi *count* di sel tersebut dengan jumlah total *case*. Misalnya, 20% adalah persentase total dari jumlah pendapatan < Rp. 250.000 dan rumah tipe 45 dan diperoleh dengan membagi 5 dengan 25.

Angka-angka di sebelah kanan dan di bawah tabel disebut marjinal. Jika berada di baris disebut marjinal baris dan jika berada dalam kolom disebut nilai marjinal kolom.

Pada gambar 2, statistik chi-kuadrat yang dihasilkan harus dibandingkan dengan titik

kritis dari distribusi teoritis chi-kuadrat untuk menentukan apakah kedua variabel benar-benar independen. Untuk itu diperlukan juga derajat kebebasan (*degree of freedom/df*) dari tabel. Derajat kebebasan untuk tabel yang terdiri dari *r* baris dan *c* kolom adalah  $(r-1) \times (c-1)$ .

Dari hasil olahan terlihat bahwa probabilitas sebuah sampel random yang berada dalam nilai chi-kuadrat yang dihasilkan adalah sedikitnya 0.00016. Probabilitas ini juga dikenal sebagai taraf signifikan pengamatan. Jika probabilitas (taraf signifikansi) yang dihasilkan cukup kecil (biasanya lebih kecil dari 0.05 atau 0.01), maka hipotesis yang mengatakan bahwa kedua variabel tersebut independen, ditolak.

**Gambar 2.** Tabulasi silang dan statistik Chi-Kuadrat

JUMLAH		Jumlah Pendapatan by TIPERUMA Tipe Rumah			Page 1 of 1	
		TIPERUMA				
		Count				
		Exp Val	tipe 21	tipe 36	tipe 45	Row
JUMLAH		Residual	1.00	2.00	3.00	Total
1.00			1	0	5	6
< Rp. 250.000			1.4	1.0	3.6	24.0%
			-.4	-1.0	1.4	
2.00			2	2	6	10
Rp. 250.000 - Rp.			2.4	1.6	6.0	40.0%
			-.4	.4	.0	
3.00			3	2	4	9
> Rp. 500.000			2.2	1.4	5.4	36.0%
			.8	.6	-1.4	
	Column		6	4	15	25
	Total		24.0%	16.0%	60.0%	100.0%

  

Chi-Square	Value	DF	Significance
Pearson	2.71296	4	.00016
Likelihood Ratio	3.60325	4	.00015
Mantel-Haenszel test for linear association	1.53655	1	.61344

Minimum Expected Frequency = .960  
 Cells with Expected Frequency < 5 = 7 of 9 (77.8%)

Number of Missing Observation: 0

Alternatif dari uji Pearson chi-kuadrat yang biasa digunakan adalah *Likelihood-ratio chi-square*. Uji ini didasarkan pada teori *maximum-likelihood* dan sering digunakan dalam analisis data kategori. Untuk ukuran sampel yang besar, pengujian dengan Pearson maupun *likelihood-ratio* akan memberikan hasil yang sama. Sedangkan uji *Mantel-Haenszel chi-square* hanya diaplikasikan untuk data dalam pengukuran ordinal.

Uji chi-kuadrat hanyalah uji independensi, sehingga hanya sedikit memberikan informasi mengenai kekuatan atau bentuk asosiasi antara dua variabel. Harga chi-kuadrat yang dihasilkan tidak hanya bergantung pada kebaikan model independensi, akan tetapi juga bergantung pada ukuran sampel. Jika ukuran sampel pada sebuah tabel ditambah maka harga chi-kuadrat juga bertambah. Jadi, besarnya harga chi-kuadrat bisa dihasilkan melalui residual yang relatif kecil untuk frekuensi harapan akan tetapi ukuran sampelnya besar.

Untuk mendapatkan pendekatan distribusi chi-kuadrat yang baik, beberapa kondisi harus dipenuhi, yaitu: data harus berupa sampel random dari distribusi multinomial dan harga pengharapan harus tidak terlalu kecil. Bila seluruh frekuensi harapan paling sedikit 5, hal ini menunjukkan kemungkinan yang sangat kuat.

## V. INTERPRETASI HASIL

Interpretasi hasil analisis tabulasi silang adalah bentuk pengujian keterhubungan antar variabel menggunakan Uji *Chi Square Pearson* dan/atau *likelihood ratio*, pengukuran derajat asosiasi antar dua variabel/faktor menggunakan koefisien kontingensi dan penentuan jenis asosiasi antar variabelnya untuk mendefinisikan jenis masing-masing variabel [Apakah variabel tak bebas (*dependent*) atau variabel bebas (*independent*)?]. Pengujian keterhubungan didasarkan pada dua jenis pengukuran yaitu pengukuran nominal dan pengukuran ordinal. Masing-masing jenis pengukuran ini akan

berbeda pada alat ukur yang digunakan untuk pengukuran.

### 5.1. Pengukuran Nominal

a. Pengukuran berdasarkan chi-kuadrat  
Pengukuran nominal yang didasarkan pada chi-kuadrat adalah hasil modifikasi dari statistik chi-kuadrat yang digunakan untuk meminimalkan pengaruh ukuran sampel dan derajat kebebasan maupun membatasi rentang harga pengukuran antara 0 s/d 1.

Koefisien Phi diperoleh dari persamaan:

$$\Phi = \sqrt{\chi^2 / N}$$

Khusus untuk tabel 2 x 2 koefisien phi yang dihasilkan sama dengan koefisien korelasi Pearson. Untuk tabel yang mempunyai dimensi (banyaknya kategori kolom atau kategori baris dalam tabel) lebih besar dari 2, maka phi yang dihasilkan bisa tidak berada di antara 0 dan 1. Hal ini dikarenakan harga chi-kuadratnya bisa lebih besar dari ukuran sampelnya. Oleh karena itu, untuk mencapai nilai antara 0 dan 1 Pearson menciptakan pengukuran sebagai berikut :

$$C = \sqrt{\chi^2 / \chi^2 + N}$$

Nilai C ini disebut sebagai Koefisien Kontingensi (*Contingency Coefficient*). Meskipun nilainya terletak antara 0 dan 1 namun pada umumnya nilai C tidak bisa mencapai limit 1. Harga maksimum koefisien ini bergantung pada banyaknya baris dan banyaknya kolom tabel. Sebagai contoh: untuk tabel berukuran 4 x 4, harga maksimum dari koefisien C adalah 0,87.

b. Nilai Cramer's V

Terdapat varian lain di perhitungan berdasarkan chi-kuadrat yang diperkenalkan oleh Cramer dan disebut nilai *Cramer's V*, yang dituliskan dalam persamaan berikut :

$$V = \sqrt{\chi^2 / N(k-1)}$$

di mana k adalah banyaknya baris atau kolom terkecil. Harga yang dihasilkan bisa

mencapai 1 untuk tabel dengan sembarang dimensi. Jika salah satu dimensi tabel adalah 2, maka  $V$  dan  $\Phi$  adalah identik.

c. Reduksi Proporsional Pada Error

Reduksi proporsional pada error (*proportional reduction in error-PRE*) adalah dasar dari berbagai pengukuran yang didasarkan pada chi-kuadrat. Dengan ukuran PRE, arti dari asosiasi menjadi lebih jelas. Pengukuran ini pada dasarnya adalah rasio dari pengukuran *error* dalam memprediksi nilai-nilai sebuah variabel yang didasarkan pada variabel itu sendiri dan pengukuran error yang sama diaplikasikan untuk memprediksi berdasarkan sebuah variabel tambahan.

Pada hasil *output*, misalnya, terlihat bahwa yang berlaku sebagai kategori modal (kategori keberhasilan dari observasi dengan proporsi terbesar -- ramalan yang terbaik dari hasil perlakuan bilamana tidak ada informasi lain yang tersedia adalah dengan menggunakan kategori modal yang ada) adalah Tipe 45 karena merupakan kategori hasil terbesar, yaitu 38% dari subyek. Estimasi probabilitas dari klasifikasi yang tidak tepat (*incorrect*) adalah 1 dikurangi probabilitas kategori modal, yaitu:

$$P(1) = 1 - 0,38 = 0,62$$

Informasi yang dihasilkan oleh besar pendapatan keluarga (bpk) dapat digunakan untuk memperbaiki aturan pengklasifikasian. Untuk masing-masing kategori bpk, kategori hasil (variabel tipe rumah) yang mempunyai frekuensi terbesar untuk kategori bpk yang bersangkutan adalah yang diprediksi. Jadi, kategori tipe 21 diprediksi untuk partisipan bpk <Rp.250.000,-, tipe 36 diprediksi oleh bpk >Rp.500.000,- dan tipe 45 diprediksi oleh bpk Rp.250.000,- - Rp. 500.000,-. Probabilitas error yang menggunakan bpk untuk memprediksi hasil adalah jumlah probabilitas seluruh sel yang tidak merupakan modus baris, jadi :

$$P(2) = 0,06 + 0,05 + 0,08 + 0,1 + 0,06 + 0,12 = 0,47$$

Lambda dari Goodman dan Kruskal yang hasilnya (variabel tipe rumah) dijadikan variabel prediksi (variabel dependen) diperoleh dari :

$$\lambda = \frac{P(1) - P(2)}{P(1)} = \frac{0,62 - 0,47}{0,62} = 0,242$$

Jadi diperoleh reduksi 24,2% pada error bilamana level besar pendapatan keluarga digunakan untuk memprediksi hasil.

Harga lambda yang dihasilkan selalu di antara 0 dan 1. Harga 0 berarti variabel independen tidak bisa memprediksi variabel dependen, sedangkan harga 1 berarti variabel independen sangat jelas menentukan kategori-kategori variabel dependen (harga sempurna ini hanya akan terjadi bila masing-masing baris mempunyai paling banyak satu sel yang tidak kosong). Jika kedua variabel independen, maka harga lambdanya 0, tetapi tidak berlaku sebaliknya.

Sebagaimana pengukuran asosiasi lainnya, lambda juga merupakan pengukuran asosiasi yang sangat spesifik. Dalam hal tertentu, lambda merefleksikan reduksi pada *error* bilamana nilai-nilai suatu variabel digunakan untuk memprediksi nilai-nilai variabel lain. Jika tipe asosiasi tidak ada, maka harga lambdanya 0, tetapi tidak berlaku sebaliknya.

Untuk tabel tertentu, dua harga lambda yang berbeda bisa dihasilkan, satu menggunakan variabel baris sebagai prediktor dan yang lain menggunakan variabel kolom sebagai prediktor.

## 5.2. Pengukuran Ordinal

Hubungan antar variabel ordinal sebenarnya dapat diuji dengan pengukuran nominal, tetapi masih perlu ditambah dengan pengukuran yang berkenaan dengan ranking. Hal yang dipertimbangkan dari hubungan yang terjadi antar variabel ordinal adalah arah hubungan dan konsep korelasi. Variabel-variabel akan mempunyai korelasi

positif jika *case-case* yang bernilai kecil dari suatu variabel cenderung mempunyai nilai yang kecil pula untuk variabel yang lain, dan *case-case* yang bernilai besar untuk suatu variabel cenderung bernilai besar pula untuk variabel lainnya. Sebaliknya, untuk variabel-variabel yang mempunyai korelasi negatif jika mempunyai nilai yang kecil pada suatu variabel, pada variabel lainnya cenderung mempunyai nilai yang besar dan sebaliknya.

Koefisien korelasi Spearman biasa digunakan untuk mengukur korelasi antar dua variabel ordinal. Untuk seluruh *case*, *value-value* dari masing-masing variabel diranking dari yang kecil hingga yang besar, dan pada ranking tersebut dihitung koefisien korelasi Pearson-nya.

Dalam tabulasi silang, untuk mengukur asosiasi linier antara variabel baris dan variabel kolom digunakan *Mantel-Haenszel Chi-Kuadrat*. Statistik yang diperoleh dengan mengkuadratkan koefisien korelasi Pearson dan hasilnya dikalikan dengan banyaknya *case* minus 1:  $\text{Mantel-Haenszel} = (\text{Koefisien korelasi Pearson})^2 * (N-1)$

Statistik *Mantel-Haenszel chi-kuadrat* mempunyai derajat kebebasan 1. Statistik ini tidak bisa digunakan untuk data nominal, meskipun SPSS selalu menampilkannya setiap kali chi-kuadrat diperlukan.

#### Pengukuran ordinal berdasarkan pasangan

Untuk tabel dari dua variabel ordinal, beberapa pengukuran asosiasi didasarkan pada perbandingan nilai-nilai dari kedua variabel untuk seluruh pasangan-pasangan *case-case* yang mungkin atau observasi-observasi yang tersedia. Kasus-kasus tersebut pertama kali harus dibandingkan untuk menentukan apakah pasangannya *concordant*, *discordant*, atau *tied*.

Pasangan kasus tersebut disebut *concordant* apabila *value-value* dari kedua variabel untuk salah satu *case* lebih tinggi (atau lebih

rendah) dibandingkan *value-value* yang berkorespondensi untuk *case* yang lain. Pasangan *case* lebih disebut *discordant* jika *value* dari salah satu variabel untuk sebuah *case* lebih besar dibanding *value* yang berkorespondensi untuk *case* yang lain, dan arahnya dibalik untuk variabel kedua. Sedangkan pasangan disebut *tied* apabila kedua *case* mempunyai nilai-nilai yang identik pada salah satu atau kedua variabel.

Jadi untuk sembarang pasangan *case* yang diberikan dengan pengukuran pada variabel X dan Y, maka pasangan tersebut bisa *concordant*, *discordant*, atau *tied* dalam satu dari tiga cara, yaitu mungkin *tied* pada X tapi tidak pada Y, mungkin *tied* pada Y namun tidak pada X atau mungkin *tied* pada kedua variabel. Jika data disusun dalam tabulasi silang, maka banyaknya pasangan yang *concordant*, *discordant* atau *tied* dapat dengan mudah dihitung, karena seluruh kemungkinan pasangan dapat ditentukan dengan baik.

Jika pasangan-pasangan tersebut lebih banyak yang *concordant*, maka asosiasinya dikatakan positif sebagaimana naiknya (atau menurunnya) rank variabel X seiring dengan menaikinya (menurunnya) rank variabel Y. Jika sebagian besar pasangan tersebut *discordant*, maka dikatakan asosiasinya negatif sebagaimana naiknya rank salah satu variabel yang diikuti dengan turunnya rank variabel yang lain. Jika pasangan *concordant* dan *discordant* berkemungkinan sama, maka dikatakan tidak ada asosiasi.

Pengukuran ordinal yang dijelaskan di sini seluruhnya mempunyai pembilang yang sama, yaitu : banyaknya pasangan concordant (P) dikurangi banyaknya pasangan discordant (Q) yang dihitung untuk semua pasangan pengamatan yang berbeda. Hasil P-Q yang dibagi dengan banyaknya pasangan merupakan ukuran yang sederhana. Ukuran ini disebut *Kendall's tau-a* yang jika tidak ada pasangan dengan *ties*, maka *range* dari harga-harga yang mungkin akan semakin sempit karena *range* ini bergantung pada

banyaknya *ties*. Hal ini dikarenakan semua observasi yang berada dalam baris yang sama adalah *tied*, demikian juga yang berada dalam kolom yang sama, dan hasil ukuran *tau-a* ini sulit sekali diinterpretasikan.

Ukuran yang digunakan untuk menormalkan P-Q dengan mempertimbangkan *ties* pada masing-masing variabel dalam sebuah pasangan secara terpisah tetapi tidak untuk *ties* pada kedua variabel dalam suatu pasangan, adalah tau-b :

$$\tau_b = \frac{P-Q}{\sqrt{(P+Q+T_x)(P+Q+T_y)}}$$

dengan  $T_x$  adalah banyaknya pasangan *tied* pada X tetapi tidak pada Y, dan  $T_y$  adalah banyaknya pasangan *tied* pada Y tetapi tidak pada X. Jika tidak ada frekuensi marginal yang berharga 0, maka tau-b dapat mencapai +1 atau -1, akan tetapi hanya untuk tabel bujursangkar.

Suatu ukuran yang dapat mencapai atau mendekati +1 atau -1 untuk tabel berukuran  $m \times n$  adalah tau-c :

$$\tau_c = \frac{2r(P-Q)}{N^2(r-1)}$$

dengan  $r$  adalah banyaknya baris terkecil atau kolom terkecil.

*Goodman and Kruskal's Gamma* berhubungan erat dengan statistik tau dan diperoleh dari :

$$G = \frac{P-Q}{P+Q}$$

Gamma dapat dianggap sebagai probabilitas bahwa pasangan random dari observasi-observasi adalah *concordant* dikurangi probabilitas bahwa pasangan tersebut *discordant*, dengan asumsi tidak ada pasangan *tied*. Harga absolut dari gamma merupakan reduksi proporsional dalam *error* antara perkiraan ranking *concordant* dan *discordant* dari masing-masing pasangan bergantung

pada kejadian mana yang paling sering dan *ranking* yang bergantung pada *toss* (pelemparan koin). Gamma akan berharga 1 jika seluruh observasi terkonsentrasi pada diagonal tabel kiri atas hingga kanan bawah. Dalam kasus independen, gamma akan berharga 0. Akan tetapi jika gamma berharga 0, belum tentu kasusnya independen, kecuali untuk tabel berukuran  $2 \times 2$ .

Dalam menghitung harga gamma tidak dibedakan variabel independen dan variabel dependen, keduanya dianggap simetris. Somers membuat koefisien *dy* (*Somer's d*) yang merupakan perluasan asimetris dari koefisien gamma. Koefisien ini dibedakan hanya dengan memasukkan banyaknya pasangan yang bukan *tied* pada variabel independen (X) pada penyebutnya.

$$dy = \frac{P-Q}{P+Q+T_y}$$

## VI. KESIMPULAN

Tabulasi silang dapat diaplikasikan di dalam perencanaan wilayah dan kota terutama di dalam menganalisis keterkaitan antara dua variabel yang mempunyai tipe data terutama dalam bentuk data kualitatif. Tabulasi silang mengolah data dan melakukan tiga tahap perhitungan:

1. Perhitungan-perhitungan statistik dalam bentuk uji-uji hipotesa yang sangat beragam akan membantu dalam menguji keterkaitan antara dua variabel.
2. Tabulasi silang juga menghasilkan koefisien-koefisien yang menunjukkan derajat hubungan antara dua variabel tersebut
3. Tabulasi silang menentukan arah hubungan yang terjadi, yaitu dengan menentukan variabel mana yang bebas dan variabel mana yang tidak bebas.

Dengan ketiga tahapan tersebut, maka tabulasi silang akan membantu di dalam menganalisis tahap selanjutnya, sebagai dasar dalam suatu pengambilan keputusan.

## VII.REFERENSI

- Dillon, William dan Matthew Goldstein.  
1984. "*Multivariate Analysis - Methods and Applications*". New York: John Wiley & Sons Inc.
- Kachigan, Sam Kash. 1986. "*Statistical Analysis - An Interdisciplinary Introduction to Univariate & Multivariate Methods*". New York: Radius Press.
- Sawitri, Dewi dan Iwan P.Kusumantoro.  
1992. "Penerapan Analisis Faktor Dalam Perencanaan Wilayah dan Kota", *Jurnal Perencanaan Wilayah dan Kota*, No5/ Triwulan III/ September 1992.
- Indratno, Imam dan Rahmat Irwinsyah.  
1997. "Modul Praktikum Analisis Tabulasi Silang", Laboratorium MAP, Jurusan Teknik Planologi UNISBA.